Matt Crawford
Amitoj Singh
Adam Walters
Jason Allen
June 15, 2011, revised July 7, 2011

ACTION PLAN FOR GCC CRB COOLING

This report explains the observations and reasoning behind the GCC load-shed action plans. The associated ITIL Problem Investigation is PBI000000000256. Load shedding is managed through the critical incident process.

The nature of the problem

GCC Computer Room B's total power is, and has always been, significantly below its designed maximum, and below the rating of the CRACs (Computer Room Air Conditioners). When at least 9 out of the 10 CRACs are operating, they should be able to remove 945 kW of heat. The room is intended to operate up to 900 kW of computing, and has been at 700 kW in recent months. It was around 600 kW in the summer of 2010.

Without short-cutting the root cause analysis, we can say that the difficulties appear to center on the condensers and the circumstances they operate under. In the worst weather conditions, a condenser may not be getting enough air to fully condense the coolant sent by the CRAC, or the air may be too hot. There are notes and photos attached to the PBI in Remedy to explain this more fully, and CD DocDB Document 4359 gives a background on our experience with cooling CRB.

Once the condensers start returning a liquid/gas mixture, effectiveness of the cooling system falls, the compressors in the CRACs run harder, but the room temperature rises only slightly until one or more CRACs fail outright, or give the "HIGH HEAD" alarm (signaling high pressure at the coolant return) and their compressors stop. Then the room temperature rises rapidly and CRAC failures accelerate.

Shedding heat load

The first critical incident involving GCC cooling in 2011 was from June 6–9. Soon afterward, several service and stakeholder representatives prepared the first draft of a staged power reduction plan to be used if needed to prevent another cooling failure if technical experts considered such an event imminent. This plan was revised with input from more stakeholders in early July. The first stage involves reducing power by about 30%, shutting down mostly older-vintage systems in racks 3000-3027. We have considered a possible second stage of power reduction by another 20 to 30%. However, we strongly

doubt that a second stage would prevent a cooling failure if the first reduction does not[1]. The final stage, of course, is shutting off all computing equipment.

Racks to power off in stage one, listed by the half-row

3003	3007–3012
3014–3020	3021–3027
(none)	(none)
3054–3055	3042
(none)	(none)
(none)	3075

Affected stakeholders

The racks listed above contain computing equipment used by the following primary stakeholders. It should be borne in mind that a count of racks shut down is not a complete measure of the impact on stakeholders. Racks contain computers of different power, and stakeholders have different amounts of computing power in other locations. This table shows the racks to be shut down in the first stage and the total racks in GCC CRB used by each stakeholder.

CDF	CMS	D0	FermiCloud	GPCF	GPFarm	GridServices
5/16	10/36	6/15	1/2	0/2	3/11	0/1

Invoking the load-shed plan

Comparison of last summer's cooling incidents to Fermilab weather records suggests that the following three factors combine to cause cooling failures:

- Outside temperature above 88°F, especially for consecutive days,
- High solar radiation to the ground,
- Low wind.

We believe that waiting until the hottest part of the day to reduce the heat load is risky. Once multiple CRAC compressors are running at 75 to 100% of capacity, the room may already be in imminent danger of cooling failure. Reducing the load mid-morning when

[1] "No plan survives first contact with the enemy!"—von Moltke the elder. When this load shed plan was put into action on July 19 and 30% reduction did not seem sufficient, ad hoc incremental shutdowns were made in order to keep some computing running to prepare results for the EPS meeting.

adverse conditions are expected is safer, but of course has a greater impact on computing throughput and a risk of a false preventive action. Invoking the plan in the morning allows orderly communication through the Service desk. System managers can then shut down their designated machines remotely. If load reduction happens in the afternoon, implementation may have to be rushed, even to the point of turning off power at the electrical panels.

Facilities will notify the Service Manager as early as possible when the former judges that conditions may require shedding heat load in GCC. Stakeholders and service providers will be notified through the Service Desk. If the load shed is needed, service providers will perform the shutdowns and a critical incident team will be assembled. Restoration of service will be decided by the critical incident team.